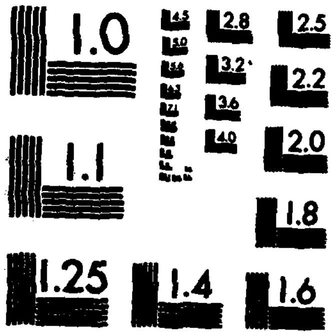
AD-A142 977 ESTIMATION OF AFEES (ARMED FORCES ENTRANCE EXAMINATION STATIONS) SURVEY WEIGHTS(U) RAND CORP SANTA MONICA CA R BUDDIN JAN 84 RAND/N-2072-MIL 1/1

UNCLASSIFIED F/G 12/1 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

(12)

A RAND NOTE

ESTIMATION OF AFEES SURVEY WEIGHTS

Richard Buddin

January 1984

N-2072-MIL

Prepared for

The Office of the Assistant Secretary of [illegible]
Manpower, Installations and Logistics

Rand

The research described in this report was sponsored by the Office of the Assistant Secretary of Defense/Manpower, Reserve Affairs and Logistics under Contract No. MDA903-83-C-0047.

Published by The Rand Corporation

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER N-2072-MIL	2. GOVT ACCESSION NO. A142977	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Estimation of AFEES Survey Weights		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard Buddin		8. CONTRACT OR GRANT NUMBER(s) MDA903-83-C-0047
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, CA. 90406-2138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Manpower, Installations and Logistics Assistant Secretary of Defense Washington, D.C. 20301		12. REPORT DATE January 1984
		13. NUMBER OF PAGES 29
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No Restrictions

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Surveys
Statistical Analysis
Enlisted Personnel
Weighting Functions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

The 1979 DoD Survey of Personnel Entering Military Service was administered to individuals signing military enlistment contracts at Armed Forces Entrance Examination Stations (AFEES). As with all surveys, some individuals in the sampled population did not respond and complete an AFEES Survey. This Note describes a methodology used to develop weights which make the respondent sample more representative of the underlying population with respect to known population demographic parameters. The weights have fairly high efficiency, so population inferences based on the weighted broaden the applications of the database. The Note also discusses appropriate uses of the weighted file and briefly reviews several recent articles on applications of survey weights to statistical analysis.

A RAND NOTE

ESTIMATION OF AFEES SURVEY WEIGHTS

Richard Buddin

January 1984

N-2072-MIL

Prepared for

The Office of the Assistant Secretary of Defense
Manpower, Installations and Logistics

Rand
1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

PREFACE

The *1979 DoD Survey of Personnel Entering Military Service* was administered at local Armed Forces Entrance Examination Stations (AFEES) to assess the motivation and background of non-prior-service military enlistees. Many researchers have been concerned about the representativeness of the survey responses, because the response rate was a modest 56 percent. This Note describes estimated survey weights that correct for differences between the respondent sample and the eligible population in terms of enlistee demographic characteristics and survey administrative differences. These weights broaden possible applications of the survey by making the respondent group more representative of the underlying population.

The Note was prepared under Department of Defense Contracts MDA-903-80-C-0652, Task Order 82-V-1, and MDA 903-83-C-0047, Task Order 83-I-2, as part of the the work of The Rand Corporation's Defense Manpower Research Center. It was sponsored by the Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics).

SUMMARY

The *1979 DoD Survey of Personnel Entering Military Service* was administered to individuals signing military enlistment contracts at local Armed Forces Entrance Examination Stations (AFEES). The survey collected detailed background and motivational information for use in research and policy decisions in the areas of accession and first-term attrition. Current applications of the survey have been limited by concern about a response bias due to the 56 percent response rate in the survey. This research examines differences between survey respondents and the eligible population and describes a procedure to develop survey weights that adjust for these differences.

Although survey information is available only for respondents, many population characteristics are known for all eligible recruits who enlisted during the prescribed survey period at each AFEES. Population and sample groups are compared across individual demographic variables such as education level, age, race, and sex. Individual refusal rates differ across these variables in several civilian surveys. Differences in response were also compared across variables which reflected possible differences in survey administrative framework. These administrative variables are AFEES, service choice, and participation in a delayed entry program. These differences could influence the availability of survey forms, the amount of encouragement for compliance, and time for compliance.

Considered separately, each factor influences response rate significantly. A log-linear model is estimated that simultaneously controls for response differences across characteristics and isolates the primary observed factors influencing response. Ultimately, most factors are insignificant after adjusting for differences in response rate by delayed entry participation and AFEES location. The log-linear procedure separates random differences in response across characteristics from systematic differences and enhances the precision of the derived survey weights.

The estimated weights remove most of the response bias due to observed differences in individual and administrative characteristics. The weights also have fairly high efficiency, so population inferences based on the weighted estimates are precisely estimated. In each wave, standard confidence intervals for population means are only about 15 percent larger for weighted as contrasted with unweighted estimates.

The weighted survey is useful primarily to derive population inferences for means and proportions. For most common regression applications, survey weights are not needed to derive unbiased and efficient parameter estimates. Some unknown response bias may persist in the survey if response is systematically related to variables which are not available for the weighting analysis.

ACKNOWLEDGMENTS

Zahava Doering of the Office of the Assistant Secretary of Defense (MI&L) offered support, encouragement, and advice for this research. Helen Hagen and John Richards of the Defense Manpower Data Center created the data sets that were used to develop the sample weights. Among my Rand colleagues, I am grateful to Robert Bell, James Press, and Karl Schutz for their statistical advice and expertise. James Hosek and Daniel Relles provided valuable suggestions on an earlier draft. Jeanne Heller provided expert editorial advice on the final draft.

CONTENTS

PREVIOUS PAGE IS BLANK

TABLES

I. INTRODUCTION

The *1979 DoD Survey of Personnel Entering Military Service* was administered to individuals signing military enlistment contracts at Armed Forces Entrance Examination Stations[1] (AFEES) and is typically referred to as the 1979 AFEES Survey. The purpose of the survey was to aid policy decisions and research in the areas of accession and first-term enlisted attrition. It contains detailed information on individual motivation and background at the time of enlistment. The survey was administered to all non-prior-service enlistees during four week periods in the spring and fall of 1979. Doering et al. (1980a, 1980b) provide a detailed description of the design, administration, and contents for the spring and fall waves.

As with all surveys, some individuals in the sampled population did not respond and complete an AFEES Survey. The response rate during the spring wave of the survey was 55.8 percent and 56.0 during the fall wave. Less than 100 percent response creates the possibility that nonrespondents may differ systematically from respondents and that inferences drawn from respondents may provide misleading indications of the behavior patterns of all enlistees. This Note describes a methodology used to develop weights which make the respondent sample more representative of the underlying population with respect to known population demographic parameters. The paper complements existing documentation of the AFEES file, and the survey weights will broaden the applications of the database. Use of the unweighted (self-weighted) survey implicitly relies on the assumption that respondents as a group are representative of the population overall. For many purposes, researchers may find it preferable to rely on the weighted survey and assume that respondents in a well-defined (i.e., age and service) group are representive of the overall population in that same group.

[1] AFEES are now called Military Enlistment Processing Stations (MEPS). Since the data base is known as the AFEES Survey, we have chosen to use the term AFEES throughout this Note instead MEPS.

The next section describes differences between the respondents and the eligible population. Survey weights are estimated that adjust for differences in response rates across various population characteristics and administrative units. The third section discusses the efficiency of estimates using these weights and how much bias the weights remove. The final section discusses appropriate uses of the weighted file and briefly reviews several recent articles on applications of survey weights to statistical analysis.

I. WEIGHTING PROCEDURE

BACKGROUND

Individual and administrative factors influence the level and characteristics of survey response. Previous survey experience has shown that response rates frequently vary with the individual characteristics of the survey population such as age, sex, and race-ethnicity (Frankel and McWilliams, 1981; Institute for Social Research, 1972; Jones et al., 1983). The administrative framework used for surveying may also influence response rates. Ideally, survey administrators would understand the purposes of the survey, encourage individuals to complete the survey, and provide adequate time for survey completion. These objectives are probably less uniformly met when the survey is administered in a variety of places by different people, as was the AFEES Survey.

Considerable information is available to assess the extent of possible individual and administrative response biases in the AFEES Survey. The eligible population consists of non-prior-service enlistees during a prescribed twenty-day working period at each AFEES. An enlistment record is generated on enlistment day that describes demographic attributes of each individual recruit. The enlistment records processed during the survey period in each AFEES provide a description of recruits in terms of education level, age, race-ethnicity, sex, service choice, and Delayed Entry Program (DEP) participation.[1] The distribution of these variables in the eligible survey population can be compared with their distribution among survey respondents to assess the dimensions of survey response.[2]

---

[1]DEP is a program that allows delays between enlistment (signing a military enlistment contract) and the actual start of active military duty. DEP is common in all services and in 1979 could last for up to twelve months. Program participation may reflect several factors including a desire to complete a school term or work commitment, waiting for a training slot opening, or taking time off before entering the military.

[2]All available comparable variables are used in the weighting analysis with two exceptions. Response rates could have been compared across states, but AFEES location identification was used because

Survey administration may vary with AFEES location, DEP, and service. Different people were responsible for the distribution and collection of surveys at the 67 AFEES where the survey was administered in fall and spring waves. It seems almost inevitable that these differences in the administration environment would produce less than uniform survey response.[3] Individuals who are entering the service directly and not participating in DEP require more processing through the AFEES and may not have equal access to the survey or time to complete the survey. Similarly, processing requirements at the AFEES may vary by service, so the response rate may vary with service choice.

Individual differences in response are expected to vary with age, sex, race-ethnicity, and educational level. A variety of psychological and sociological explanations have been offered to explain why differences in these types of demographic characteristics influence survey response. Actual enlistment demographic information can be compared with survey demographic information to determine whether any of these factors influenced response in the AFEES Survey.

Separate weighting procedures were appropriate for the fall and spring waves of the survey, because the survey designers believed that recruit backgrounds and enlistment motivations varied between the spring and fall. Common weights for both waves would require the restrictive assumption that respondents in one wave differ from respondents in the other only in terms of a few factors which are available on the eligible population in each period. Most of the background and motivational variables like family income, reasons for enlisting, and military job availability are not available on enlistment records and could not be used for weighting. These background and motivation variables are probably correlated with available weighting data on individual

---

surveys were actually administered at the AFEES stations. Marital status at enlistment was available for respondents and the eligible population, but marital status for the eligible population was frequently missing, e.g., marital status is missing from enlistment records for 40 and 64 percent fo the spring and fall waves, respectively. With such large missing categories, I was unable to use marital status as a weighting variable.

[3]Doering et al. (1980a, 1980b) noted that some AFEES "did not always follow instructions for collecting data and identifying respondents."

demographic characteristics. As a result, a single set of weights for both waves might inappropriately distort differences between the background and motivation of spring versus fall enlistees. Separate spring and fall weights were developed to make sampled respondents more representative of the surveyed population in the respective time frame. Comparisons of the spring and fall samples are reserved for analysis and not addressed here.

PATTERNS OF NONRESPONSE

If response were random, we would expect response rates for most AFEES to cluster around the 56 percent average response rate for each wave. Table 1 indicates a substantial difference in the survey response rate across AFEES.[4] The distribution of AFEES around the mean response

Table 1

DISTRIBUTION OF RESPONSE RATES ACROSS AFEES

Response Rate	Percent of AFEES	
	Spring Wave	Fall Wave
<30	6.3	9.1
30-39	9.5	12.1
40-49	14.3	12.1
50-59	22.2	7.6
60-69	14.3	21.2
70-79	15.9	12.1
80-89	14.3	21.2
90+	3.2	4.5
(n)	63	65
χ^2	2435.6	1897.1

[4]Response rates by AFEES are reported in appendix Table A-1. AFEES proportions in the survey and eligible population are reported in appendix Tables A-2 and A-3. The survey was not administered in Syracuse, New York in either wave. The survey was also not administered in Manchester, New Hampshire and Baltimore, Maryland in the spring wave. Responses from the Los Angeles and San Diego AFEES are combined, because separate enlistment records were not maintained.

rate in the spring wave is symmetric, but fairly large proportions are in the extremes, e.g., 15.8 percent of the AFEES had response rates less than 40 percent and 17.5 percent had rates greater than 80 percent. In the fall wave, the distribution of response rates across AFEES is bimodal with few stations reporting response in the middle 50-59 percent range. Larger percentages of stations had high response rates in the fall than in the spring, but larger percentages of stations had low response rates in the fall than the spring. These differences in response rate across AFEES suggest that the survey was not administered uniformly and/or refusals varied systematically with location. Large differences in response rate by AFEES enhance the likelihood that the respondent group may be unrepresentative of the survey eligible group in other respects.

Tables 2 and 3 describe population and survey differences in the distribution of individual and administrative characteristics for the spring and fall waves. The chi-square tests indicate that the distribution of survey respondents differs significantly from that of the eligible population for most characteristics in both waves. For example, the chi-square statistic for education level means that we cannot accept the null hypothesis of no difference between the education distribution in the survey sample and the population. In both waves, high school graduates are more likely respondents than non-high-school graduates. Response rates increase monotonically with age in the fall wave, whereas response is highest for the youngest and oldest groups of recruits in the spring. Whites are slightly more likely to complete surveys than nonwhites. DEP participants have lower response rates in the spring wave than recruits who are entering the service directly, but the pattern is reversed for the fall wave. Female recruits are slightly more likely to complete surveys than males. Finally, response varies substantially with service. Navy recruits have lower response rates than those from the other services in both waves, with the ordering of response in other services depending on wave.

Response differences reported in Tables 1 through 3 suggest that inferences drawn from the respondent group may *not* apply to the survey population. Suppose, for example, we wanted to compare the average wage

Table 2

SURVEY AND POPULATION CHARACTERISTICS IN SPRING WAVE

Characteristic	% Survey Respondents	% Population Eligible	Response Rate	X^2
Education Level				14.4
HS grad or beyond	65.2	63.7	57.1	
Cert of gen educ develop	6.4	6.7	53.2	
Not HS grad	28.4	29.6	53.5	
Enlistment Age				37.4
<18	18.9	18.6	56.7	
18	29.2	29.7	54.7	
19	16.9	18.5	51.0	
>19	35.0	33.2	58.9	
Race				1.7
White	69.5	69.0	56.1	
Nonwhite	30.5	31.0	54.9	
DEP Participation				148.9
Participant	84.4	87.7	53.7	
Non-participant	15.6	12.3	70.7	
Sex				0.8
Male	79.3	79.6	55.5	
Female	20.7	20.4	56.5	
Service				61.0
Army	44.9	44.8	55.9	
Navy	20.4	22.7	50.1	
Air Force	21.7	21.5	56.2	
Marines	12.4	11.0	62.9	

of enlistees in the AFEES Survey with the average wage of nonenlistees from a civilian youth survey. If wage is positively related to age, then the average wage estimated from the AFEES fall wave would overstate the true wage of individuals enlisting in fall 1979, because older recruits had higher response rates than 17 and 18 year old recruits. This simple bias can be corrected by taking a weighted average of AFEES wages by age group where the weight in each group equals the inverse of the sampling proportion in each group. If respondents in each age group were a random sample of enlistees in that group, then the weighted estimate of enlistee wages would be an unbiased estimate of the population parameter.

Sample weighting is not, of course, an ideal substitute for complete response. Returning to our example, suppose that wages were related to student status at enlistment as well as age. If student status did not affect response, then the age weights proposed above

Table 3

SURVEY AND POPULATION CHARACTERISTICS IN FALL WAVE

Characteristic	% Survey Respondents	% Population Eligible	Response Rate	χ^2
Education Level				1106.9
HS grad or beyond	64.6	54.4	66.4	
Cert of gen educ develop	7.4	4.9	84.5	
Not HS grad	27.9	40.5	38.5	
Enlistment Age				322.2
<18	21.8	24.6	49.6	
18	26.4	29.0	50.9	
19	15.7	16.9	52.0	
>19	36.0	29.5	68.3	
Race				2.4
White	68.8	68.2	56.4	
Nonwhite	31.2	31.8	52.2	
DEP Participation				57.7
Participant	89.7	87.7	57.2	
Non-participant	10.3	12.3	46.9	
Sex				10.1
Male	80.2	81.2	55.3	
Female	19.8	18.8	58.9	
Service				44.8
Army	46.6	44.9	58.1	
Navy	21.2	23.3	50.9	
Air Force	20.4	19.9	57.4	
Marines	11.4	11.9	53.6	

would be sufficient for unbiased wage estimates. Alternatively, if student status affects response after controlling for age, then the estimated average wage would be biased. Unfortunately, differences in response by student status (and numerous other variables) cannot be examined, because the variable is available from the survey and not from service enlistment records. (Ironically, the weighting "problem" is eased if all survey variables were available from enlistment records, but then analysis could proceed directly from the enlistment records, and the survey would be redundant.) Weights based on observed differences in sample and population characteristics will make the respondent group more representative of the population, but biases may remain for some applications if response is not random within each weighting class.

CHOOSING WEIGHTING CLASSES

For many situations, response weights are chosen that exhaust comparable survey and population information. A matrix is constructed which has as dimensions the number of comparable characteristics and each cell entry is the ratio of the population cell count (for example, by age, race, and sex) to the survey respondent cell count. These inverse sample weights are applied to the survey records with corresponding characteristics to control for response differences. Inverse sample weights effectively force the weighted observed cell counts to equal the expected cell counts based on the population, so chi-squares comparing the distribution of weighted sample characteristics with those of the population are zero.

AFEES Survey weights based on full-interactions of AFEES, education level, age, race, DEP participation, sex, and service would rely on a weighting matrix for each wave with over 24,000 elements. Inverse sample weights based on the sample and eligible population counts in each cell would eliminate the response bias for these observed characteristics, but these weights would not be estimated with much precision. With full classification, individual cell counts are frequently small (if not zero) for both respondent and eligible population groups in each wave. The eligible population count for the spring (fall) wave was 26,452 (27,831), and the spring (fall) survey sample contained 14,751 (15,573) observations. This complete information method would minimize the chi-square, but the variance in cell weights implies that many cell weights will misrepresent the "true" response rate from the appropriate underlying population. The high variance of weights based on small cells would also diminish the efficiency of estimates from the weighted survey.

The precision of individual cell weights can be enhanced by multivariate statistical procedures that reduce weighting classes and/or class categories. Weights based on some subset of variables may explain virtually all the differences in sample response rate. For example, response rates may not vary significantly with education level after controlling for other characteristics. Similarly, 17- and 18-year-olds may have similar response rates, but these rates may be different than those of older enlistees. Elimination of cells from the full-interaction matrix based on response patterns will improve the precision of the remaining cell weights, and the remaining cells will on average each contain more observations.

A multivariate analysis of the full-interaction model is hampered by the fact that most cells are empty--there are over 24,000 possible cells and only about 15,000 survey observations for each wave. Many of the empty cells are structurally zero in that the eligible population count is zero. Because of this empty cell problem, multivariate analysis of response differences is based on a two step procedure. As a first step, the main factors affecting response are identified from least squares regressions for each AFEES.[5] Individual response/nonresponse is estimated as a function of education, age, race, DEP participation, sex, and service. At this stage of the analysis, variables enter the estimation on a first-order level, without interaction with other included variables. When controlling for all factors simultaneously, age, DEP participation, and service are the primary factors influencing response across the majority of AFEES, and

[5]AFEES were chosen as the prelimenary analysis unit because of the large response differences across stations. Subsequent analysis examines whether AFEES level response differences persist after controlling for differences in individual characteristics which influence the refusal rate.

education, race, and sex are not important. These results suggest that a weighting procedure based on AFEES, age, DEP, and service will correct for most of the observed differences between the respondent sample and the eligible population.

After dropping education, race, and sex, the revised weighting matrix has approximately 2100 cells, but the small cell problem remains. The average cell counts are about 7 sample observations per cell and 12 population observations per cell. Some cells are much smaller than average because some AFEES have few enlistments, because the Marines are a relatively small proportion of total enlistments, and because most recruits are DEP participants. In addition, some variable interactions are probably unnecessary to explain the observed patterns of response. For example, the age or service patterns of response may not vary across AFEES.

The second step of the response analysis examines variable interactions in the revised weighting matrix. A contingency table is constructed where table entries correspond to survey and eligible population counts by AFEES, age, DEP, and service. The table is analyzed with a log-linear probability model where the log of cell counts is estimated as a function of dummy variables and interactions for all AFEES, age, DEP, service and survey/population group combinations (Bishop et al., 1975; Nerlove and Press, 1973). As a computational device, AFEES are grouped into census regions, and separate specifications are estimated within each region. Inverse sample weights are computed from the fitted values of respective population and respondent cell counts for each AFEES, age, DEP, and service classification. This procedure improved the precision of the weights in two ways. First, the small cell problem is mitigated because the model uses information from similar types of individuals in each region to estimate a sample weight. Second, insignificant interactions are deleted, and the implicit cell count of remaining cells is increased.

Tables 4 and 5 describe the factors and interactions required to explain differences in the individual and administrative characteristics of survey respondents relative to those of the eligible population. Effects were screened based on their marginal and partial contribution

Table 4

FACTORS IN SPRING WAVE RESPONSE BY CENSUS REGION

Variable	Census Region[a]									
	1	2	3	4	5	6	7	8	9	10
Age					x					
Service	x		x		x					x
DEP	x	x	x	x	x	x	x		x	x
AFEES	x	x	x	x	x	x	x	x	x	
Age*Service					x					
Age*DEP					x					
Age*AFEES					x					
Service*DEP			x		x					x
Service*AFEES			x		x					
DEP*AFEES		x	x		x		x		x	
Age*Service*DEP										
Age*Service*AFEES										
Age*DEP*AFEES										
Service*DEP*AFEES										
Age*Service*DEP*AFEES										

[a] The standard census groups are 1-New England, 2-Middle Atlantic, 3-East North Central, 4-West North Central, 5-South Atlantic, 6-East South Central, 7-West South Central, 8-Mountain, 9-Pacific, 10-Other. The only surveyed AFEES in the "other" region is San Juan, so all AFEES level interactions are not applicable.

to the chi-square of the fitted model in each region. Effects were deleted until a parsimonious set remained which explained virtually all of the differences in response within the region.[6] Individual weights were based on the fitted values of the models. Most of the interaction terms are insignificant in all models, so individual weights are estimated more precisely than in the initial fully interacted age, service, DEP, and AFEES model.

[6] The screening procedure is recommended by Morton Brown, "Screening Effects in Multidimensional Contingency Tables," *Applied Statistics*, Vol. 25, No. 1, pages 37-46.

The tables reveal that the most consistent factors in explaining the response pattern are AFEES and DEP. The first-order effects of these variables enter in almost every model estimated, and the most frequent second-order interaction employed is the AFEES-DEP interaction. This pattern suggests that much of the observed nonrandom response pattern was related to survey administration and not to differences in the demographic characteristics of the surveyed population. After other factors are controlled, age patterns in response are insignificant in all but one model. Response differences by service are significant in about half of the models estimated.

The weights were estimated for observations with nonmissing AFEES, age, DEP, and service. The missing cases are less than 2 percent for each wave. "Best guess" weights are awarded to these cases, although some researchers may prefer deletion of the cases from their analysis.[7] When AFEES is nonmissing, the modal value of the missing characteristic is imputed, and the corresponding weight is applied. When AFEES is missing, which occurs for about 1 percent of total cases, the assigned weight is based on the average weight of individuals with similar age, DEP, and service characteristics in the same wave. As a final step, the weights are normalized so the sum of the sample weights equals the size of the eligible population.

[7] The set of nonmissing survey cases are weighted to represent the survey eligible population. A complete set of weights is available, if values are assumed for missing characteristics. These imputed weights make the complete weighted survey slightly less representative of the population than the set of surveys with nonmissing AFEES, age, DEP, and service. Individual researchers can choose whether they prefer the imputed weights to a slightly smaller analysis file.

Table 5

FACTORS IN FALL WAVE RESPONSE BY CENSUS REGION

Variable	Census Region[a]									
	1	2	3	4	5	6	7	8	9	10
Age										
Service	x	x	x		x					x
DEP		x	x	x	x	x	x		x	x
AFEES	x	x	x	x	x	x	x	x	x	
Age*Service										
Age*DEP										
Age*AFEES										
Service*DEP		x	x		x					x
Service*AFEES	x	x	x		x					
DEP*AFEES		x	x		x		x		x	
Age*Service*DEP										
Age*Service*AFEES										
Age*DEP*AFEES			x							
Service*DEP*AFEES										
Age*Service*DEP*AFEES										

[a] The standard census groups are 1-New England, 2-Middle Atlantic, 3-East North Central, 4-West North Central, 5-South Atlantic, 6-East South Central, 7-West South Central, 8-Mountain, 9-Pacific, 10-Other. The only surveyed AFEES in the "other" region is San Juan, so all AFEES level interactions are not applicable.

III. PROPERTIES OF SURVEY WEIGHTS

RELATIVE PRECISION OF POPULATION INFERENCES

Survey nonresponse reduces the accuracy of population inferences from survey data. The weights developed Section II substantially reduce the bias in population estimates based on the AFEES Survey, but the weighting procedure does increase the variance of those estimates relative to estimates based on unweighted data. If weighted estimates are not very efficient, however, the estimates are not very informative about the underlying population parameters.

Consider a vector of observations Y where the various components are independent with common mean (μ) and variance (σ^2). The expectations of the weighted and unweighted mean are both equal to μ, but weighting does alter the variance of the estimated sample mean. The variance of the unweighted mean is σ^2/n, and the variance of the weighted mean is σ^2w'w/(w'1)**2, where w is a vector of weights. A useful measure of the efficiency of the weighted means compares the ratio of the standard error of the unweighted and weighted means, i.e., w'1/(nw'w)**.5. The ratio is independent of σ^2 and equals the upper bound of one when all weights are equal. High efficiencies imply that inferences based on the weighted estimates are precisely estimated.

Tables 6 and 7 report the efficiencies of the survey weights estimated in Section II. The overall efficiency of the spring and fall weights is about 85 percent, which implies that standard confidence intervals for population means are only about 15 percent larger for weighted than for unweighted estimates. Within most regions, the estimator efficiency is above 90 percent. The AFEES specific efficiencies reported in Table 7 are also quite high for virtually all AFEES in each wave. Only two spring and three fall AFEES have estimator efficiencies less than 85 percent. Over two-thirds of the AFEES level efficiencies exceed .98, which indicates that weights within these AFEES have virtually no variance. The high AFEES level efficiencies as compared with the overall efficiency reveals that most of the variance in survey weights is across AFEES and not within AFEES. The fairly high

Table 6

EFFICIENCIES OF WEIGHTED MEANS
BY REGION AND WAVE

Region	Spring Wave	Fall Wave
New England	0.975	0.785
Middle Atlantic	0.883	0.941
East North Central	0.942	0.901
West North Central	0.956	0.960
South Atlantic	0.724	0.768
East South Central	0.926	0.922
West South Central	0.951	0.979
Mountain	0.962	0.984
Pacific	0.891	0.916
Other	0.936	0.792
Overall	0.856	0.852

Note: Efficiencies are defined as the ratio of the standard errors of the unweighted estimator of the population mean to the weighted estimator of the population mean. The computed efficiency is upper bounded by one when when all weights are equal.

Observations with missing AFEES are included in the other category.

estimator efficiencies at the AFEES and aggregate levels indicate that weighted estimates cost little precision.

RESPONSE BIAS REDUCTION FROM WEIGHTING

How well do weighted survey estimates compare with observed population characteristics? The survey weights estimated in Section II adjust the sample of respondents to correspond to the eligible population in terms of a set of observed population characteristics. Classical weights based on inverse sampling probabilities in the fully saturated model would exactly adjust the corresponding cell probabilities. This classical approach was abandoned because the cell sizes in the fully classified model were too small. The log-linear

Table 7

EFFICIENCIES OF WEIGHTED MEANS
BY AFEES AND WAVE

AFEES	Spring Wave	Fall Wave
Portland, ME	0.989	0.989
Manchester, NH	na	0.996
Boston, MA	0.988	0.936
Springfield, MA	0.989	0.903
New Haven, CT	0.991	0.929
Albany, NY	0.999	0.994
Brooklyn, NY	0.881	0.988
Newark, NJ	0.988	0.941
Philadelphia, PA	0.957	0.974
Syracuse, NY	na	na
Buffalo, NY	0.992	0.977
Wilkes Barre, PA	0.999	0.984
Harrisburg, PA	0.999	0.991
Pittsburgh, PA	0.996	0.920
Baltimore, MD	na	0.814
Richmond, VA	0.964	0.979
Beckley, WV	0.986	0.991
Knoxville, TN	0.993	0.999
Nashville, TN	0.995	0.999
Louisville, KY	0.994	0.999
Cincinnati, OH	0.975	0.984
Columbus, OH	0.990	0.985
Cleveland, OH	0.968	0.913
Detroit, MI	0.985	0.962
Milwaukee, WI	0.958	0.979
Chicago, IL	0.873	0.897
Indianapolis, IN	0.972	0.969
St. Louis, MO	0.997	0.999
Memphis, TN	0.993	0.999
Jackson, MS	0.995	0.999
New Orleans, LA	0.991	0.999
Montgomery, AL	0.991	0.998
Atlanta, GA	0.724	0.701
Fort Jackson, SC	0.547	0.978
Jacksonville, FL	0.979	0.997
Miami, FL	0.918	0.991
Charlotte, NC	0.948	0.996
Raleigh, NC	0.945	0.940
Shreveport, LA	0.998	0.999
Dallas, TX	0.992	0.998
Houston, TX	0.999	0.999

San Antonio, TX	0.921	0.999
Oklahoma City, OK	0.985	0.982
Amarillo, TX	0.999	0.977
Little Rock, AR	0.995	0.999
Kansas City, MO	0.998	1.000
Des Moines, IA	0.998	0.998
Minneapolis, MN	0.999	0.999
Fargo, ND	0.997	0.999
Sioux Falls, SD	0.997	0.999
Omaha, NE	0.998	0.999
Denver, CO	1.000	1.000
Albuquerque, NM	1.000	1.000
El Paso, TX	0.996	0.997
Phoenix, AZ	1.000	1.000
Salt Lake City, UT	1.000	1.000
Butte, MT	1.000	1.000
Spokane, WA	0.996	0.999
Boise, ID	1.000	1.000
Seattle, WA	0.991	0.999
Portland, OR	0.839	0.899
Oakland, CA	0.997	0.991
Fresno, CA	0.871	0.998
Los Angeles, CA	0.873	0.998
Honolulu, HI	0.997	0.999
San Juan, PR	0.940	0.639
Overall	0.856	0.852

Note: Efficiencies are defined as the ratio of the standard errors of the unweighted estimator of the population mean to the weighted estimator of the population mean. The computed efficiency is upper bounded by one when when all weights are equal.

regression model reduced the variance of estimated sample weights but some statistically insignificant bias remains. By construction, the weighted cell count within AFEES, service, age, and DEP subgroups for each wave are insignificantly different from population counts. Tables 8 through 10 show how well the weighted survey replicates population characteristics in terms of region, education level, enlistment age, race, DEP participation, sex, and service.

The most dramatic change between the weighted and unweighted survey distributions occurs with respect to location. Table 8 shows that the distribution of weighted responses across regions is quite close to the population distribution. The substantial change of weighting on the

Table 8

EFFECT OF WEIGHTING ON ESTIMATED DISTRIBUTION BY WAVE AND REGION

Region	% Unweighted Survey	% Weighted Survey	% Population Eligible
Spring Wave			
New England	5.59	4.52	4.50
Middle Atlantic	12.65	15.68	16.15
East North Central	17.31	15.04	14.83
West North Central	9.26	7.70	7.88
South Atlantic	13.40	16.46	16.51
East South Central	7.83	8.72	8.73
West South Central	13.89	10.55	10.51
Mountain	5.86	6.48	6.63
Pacific	12.20	12.01	11.92
Other	1.61	2.42	2.26
Fall Wave			
New England	6.72	4.12	5.40
Middle Atlantic	14.32	14.14	14.78
East North Central	16.12	19.15	17.88
West North Central	7.41	6.85	7.24
South Atlantic	15.10	19.90	19.20
East South Central	7.66	7.98	7.55
West South Central	12.18	8.88	9.06
Mountain	6.43	5.31	5.00
Pacific	13.54	12.10	11.96
Other	1.55	0.37	0.51

location distribution reflects the large cross-AFEES differences in response rates, and the fact the AFEES was a key variable in each log-linear specification. The weighted survey does a better job than the unweighted survey of replicating the population distribution by region, because the weights substantially reduce the response bias within each AFEES. Appendix Tables A.2 and A.3 show the effect of weighting on the distribution of responses across AFEES for the spring and fall waves, respectively.

For characteristics other than location, the bias reductions are smaller because response initially varied less systematically with these factors. Table 9 shows how weighting changes the distribution of survey attributes by charactericstics observable for the enlistment population in the spring of 1979. The weighting adjustment closes 95 percent of

Table 9

EFFECT OF WEIGHTING ON ESTIMATED DISTRIBUTION
OF POPULATION CHARACTERISTICS IN SPRING WAVE

Characteristic	% Unweighted Survey	% Weighted Survey	% Population Eligible
Education Level			
HS grad or beyond	65.2	67.5	63.7
Cert of gen educ develop	6.4	6.3	6.7
Not HS grad	28.4	26.2	29.6
Enlistment Age			
<18	18.9	19.2	18.6
18	29.2	29.6	29.7
19	16.9	16.9	18.5
>19	35.0	34.3	33.2
Race			
White	69.5	69.4	69.0
Nonwhite	30.5	30.6	31.0
DEP Participation			
Participant	84.4	87.9	87.7
Non-participant	15.6	12.1	12.3
Sex			
Male	79.3	79.6	79.6
Female	20.7	20.4	20.4
Service			
Army	44.9	45.6	44.8
Navy	20.4	21.2	22.7
Air Force	21.7	21.1	21.5
Marines	12.4	11.5	11.0

the gap between the survey and population percentages of DEP participants. More modest improvements occur in the distributions for age, and service where the initial differences in response rates are less pronounced. Since differences in response by sex and race are inconsequential, the small effect of weighting on these distributions was expected. The weights make the education distribution slightly worse, but spring differences in response by education are not large.

Table 10 reveals how weighting of the fall wave reduces the gap between survey and population distributions for available variables. The largest gains occur for education level, where response rates among non-high-school graduates was 38 percent as compared with 66 percent for graduates. The survey weights, though based largely on AFEES and DEP, have the desired effect of substantially closing the gap between population and survey percentages by education level. Smaller improvements occur in the distribution by age, DEP, and service. Weighting has little impact on the race and sex distributions, where survey response was virtually the same across categories.

Table 10

EFFECT OF WEIGHTING ON ESTIMATED DISTRIBUTION
OF POPULATION CHARACTERISTICS IN FALL WAVE

Characteristic	% Unweighted Survey	% Weighted Survey	% Population Eligible
Education Level			
HS grad or beyond	64.6	59.9	54.4
Cert of gen educ develop	7.4	6.4	4.9
Not HS grad	27.9	33.6	40.5
Enlistment Age			
<18	21.8	22.8	24.6
18	26.4	27.0	29.0
19	15.7	15.5	16.9
>19	36.0	34.7	29.5
Race			
White	68.8	69.5	68.2
Nonwhite	31.2	20.4	31.8
DEP Participation			
Participant	89.7	89.3	87.7
Non-participant	10.3	10.7	12.3
Sex			
Male	80.2	80.3	81.2
Female	19.8	10.7	18.8
Service			
Army	46.6	45.0	44.9
Navy	21.2	22.6	23.3
Air Force	20.4	20.1	19.9
Marines	11.4	11.7	11.9

IV. APPLICATIONS OF THE SURVEY WEIGHTS

Most computations of AFEES Survey means, proportions, and cross-tabulations should rely on the weighted database. Population inferences from unweighted data are implicitly based on the assumption that respondent observations are a random sample of the population. In fact, response varies systematically with several observed population characteristics, and the weighting procedure controls for these differences. Weighting will remove response bias, if respondents within a weighting group are a random sample of that population group. Some bias may remain due to unobserved factors that influence response and cannot be controlled in the weighting procedure. Nonetheless, population inferences based on weighted estimates rely on a weaker assumption of respondent representativeness than inferences from unweighted estimates.

Regression application of the AFEES Survey weights depends on the model chosen. Weighted regression is inappropriate for the standard linear regression specifications where the behavioral coefficients are homogeneous throughout the population (DuMouchel and Duncan, 1983; Holt et al., 1980; Porter, 1973; and Smith, 1976). For the standard model, the least squares parameter estimates are best linear unbiased estimates as long as the stochastic disturbance terms have zero mean, common variance (homoskedasticity), independence (nonautoregression), and nonstochastic explanatory variables. The sampling rate within a stratum does not influence these estimates, because the model is invariate across the population.

In some situations, weighting variables may enter the standard linear regression specification as explanatory variables, but least squares estimates are still preferred over weighted regression estimates. The means of many dependent variables may vary across weighting classes, and the correctly specified model will include dummy variables to control for differences across strata. Researchers can test for nonhomogeneous coefficients across strata by interactions between suspect explanatory variables and strata dummies. In the extreme, completely separate regression specifications can be run for different strata, when strata size is sufficient.

Survey weights are important for two types of regression applications. First, weights are used to derived estimates of random coefficient regression models (Holt et al.; 1980, Porter, 1973). Second, simple least squares estimates are inappropriate when the dependent variable is a selection or weighting variable. Manski and Lerman (1977) derived the appropriate estimators for this "choice-based" case. Researchers who use the AFEES Survey to estimate models of DEP participation or service choice must explicitly deal with the nonrandom response patterns in these variables.

APPENDIX

Table A-1

SURVEY RESPONSE RATES BY AFEES AND WAVE

AFEES	Spring Wave	Fall Wave
Portland, ME	0.540	0.612
Manchester, NH	na	0.624
Boston, MA	0.811	0.885
Springfield, MA	0.652	0.650
New Haven, CT	0.538	0.368
Albany, NY	0.607	0.505
Brooklyn, NY	0.316	0.598
Newark, NJ	0.496	0.604
Philadelphia, PA	0.238	0.269
Syracuse, NY	na	na
Buffalo, NY	0.740	0.651
Wilkes Barre, PA	0.459	0.714
Harrisburg, PA	0.814	0.812
Pittsburgh, PA	0.466	0.418
Baltimore, MD	na	0.345
Richmond, VA	0.435	0.287
Beckley, WV	0.869	0.811
Knoxville, TN	0.460	0.811
Nashville, TN	0.633	0.221
Louisville, KY	0.752	0.890
Cincinnati, OH	0.547	0.781
Columbus, OH	0.810	0.825
Cleveland, OH	0.766	0.602
Detroit, MI	0.562	0.323
Milwaukee, WI	0.552	0.471
Chicago, IL	0.715	0.417
Indianapolis, IN	0.598	0.480
St. Louis, MO	0.881	0.732
Memphis, TN	0.519	0.569
Jackson, MS	0.829	0.603
New Orleans, LA	0.478	0.827
Montgomery, AL	0.294	0.327
Atlanta, GA	0.462	0.110
Fort Jackson, SC	0.271	0.302
Jacksonville, FL	0.809	0.914
Miami, FL	0.244	0.440
Charlotte, NC	0.315	0.657
Raleigh, NC	0.434	0.490
Shreveport, LA	0.935	0.705

Dallas, TX	0.607	0.683
Houston, TX	0.703	0.773
San Antonio, TX	1.217	0.965
Oklahoma City, OK	0.728	0.588
Amarillo, TX	0.523	0.617
Little Rock, AR	0.707	0.597
Kansas City, MO	0.501	0.347
Des Moines, IA	0.334	0.230
Minneapolis, MN	0.566	0.463
Fargo, ND	0.731	0.677
Sioux Falls, SD	0.752	0.869
Omaha, NE	0.875	0.741
Denver, CO	0.363	0.879
Albuquerque, NM	0.444	0.417
El Paso, TX	0.662	0.763
Phoenix, AZ	0.584	0.644
Salt Lake City, UT	0.787	0.814
Butte, MT	0.690	0.662
Spokane, WA	0.673	0.765
Boise, ID	0.642	0.810
Seattle, WA	0.544	0.826
Portland, OR	0.355	0.339
Oakland, CA	0.597	0.327
Fresno, CA	0.890	0.841
Los Angeles, CA	0.544	0.803
Honolulu, HI	0.630	0.671
San Juan, PR	0.398	1.704

Note: The response rate is defined as the number of sample observations in each category as a proportion of the population eligible. The eligibility criteria are based on survey dates at each AFEES. Recorded enlistments are actually less than survey responses in two cases. These cases presumably reflect either errors in recording the appropriate number of enlistments or survey administration for more days than reported. The estimated weights are implicitly based on the assumption that reported eligibility for each category is at least proportional to the "true" eligible population.

Table A-2

EFFECT OF WEIGHTING ON SAMPLE DISTRIBUTION BY AFEES
FOR THE SPRING WAVE

AFEES	% Unweighted Survey	% Weighted Survey	% Population Eligible
Portland, ME	0.67	0.71	0.69
Manchester, NH	na		
Boston, MA	3.18	2.22	2.18
Springfield, MA	0.96	0.82	0.82
New Haven, CT	0.75	0.75	0.78
Albany, NY	0.92	0.82	0.84
Brooklyn, NY	2.90	5.14	5.11
Newark, NJ	2.16	2.43	2.43
Philadelphia, PA	1.07	2.11	2.51
Syracuse, NY	na		
Buffalo, NY	1.78	1.33	1.34
Wilkes Barre, PA	0.79	0.92	0.97
Harrisburg, PA	1.30	0.89	0.89
Pittsburgh, PA	1.69	2.00	2.02
Baltimore, MD	na		
Richmond, VA	1.84	2.39	2.36
Beckley, WV	0.98	0.63	0.63
Knoxville, TN	0.94	1.11	1.14
Nashville, TN	1.19	1.07	1.05
Louisville, KY	1.70	1.30	1.26
Cincinnati, OH	1.25	1.25	1.27
Columbus, OH	1.62	1.11	1.11
Cleveland, OH	3.25	2.40	2.36
Detroit, MI	3.64	3.67	3.61
Milwaukee, WI	1.34	1.31	1.36
Chicago, IL	4.37	3.56	3.40
Indianapolis, IN	1.81	1.70	1.69
St. Louis, MO	3.58	2.28	2.26
Memphis, TN	1.15	1.22	1.23
Jackson, MS	1.08	0.74	0.72
New Orleans, LA	1.40	1.63	1.63
Montgomery, AL	1.74	3.24	3.29
Atlanta, GA	2.47	3.12	2.98
Fort Jackson, SC	0.98	1.98	2.03
Jacksonville, FL	3.80	2.66	2.61
Miami, FL	1.24	2.78	2.82
Charlotte, NC	0.84	1.40	1.48
Raleigh, NC	1.21	1.45	1.55
Shreveport, LA	1.18	0.70	0.70
Dallas, TX	2.41	2.23	2.21
Houston, TX	2.12	1.73	1.68

San Antonio, TX	3.53	1.70	1.61
Oklahoma City, OK	1.05	0.78	0.80
Amarillo, TX	0.30	0.29	0.32
Little Rock, AR	1.09	0.85	0.86
Kansas City, MO	1.92	2.11	2.13
Des Moines, IA	0.49	0.74	0.82
Minneapolis, MN	1.35	1.29	1.33
Fargo, ND	0.40	0.29	0.31
Sioux Falls, SD	0.53	0.37	0.39
Omaha, NE	0.95	0.59	0.60
Denver, CO	2.08	3.12	3.19
Albuquerque, NM	0.46	0.52	0.57
El Paso, TX	0.77	0.61	0.65
Phoenix, AZ	1.64	1.55	1.56
Salt Lake City, UT	0.77	0.55	0.55
Butte, MT	0.46	0.35	0.37
Spokane, WA	0.63	0.50	0.52
Boise, ID	0.42	0.35	0.37
Seattle, WA	0.94	0.93	0.97
Portland, OR	1.05	1.64	1.64
Oakland, CA	3.83	3.64	3.57
Fresno, CA	1.43	0.93	0.89
Los Angeles, CA	3.58	3.71	3.67
Honolulu, HI	0.71	0.62	0.63
San Juan, PR	1.61	2.42	2.25

Table A-3

EFFECT OF WEIGHTING ON SAMPLE DISTRIBUTION BY AFEES
FOR THE FALL WAVE

AFEES	% Unweighted Survey	% Weighted Survey	% Population Eligible
Portland, ME	0.89	0.91	0.81
Manchester, NH	0.56	0.52	0.50
Boston, MA	3.72	0.97	2.35
Springfield, MA	0.93	0.88	0.80
New Haven, CT	0.61	0.82	0.92
Albany, NY	0.89	1.10	0.98
Brooklyn, NY	3.98	3.95	3.72
Newark, NJ	2.80	2.75	2.59
Philadelphia, PA	1.18	1.94	2.45
Syracuse, NY	na		
Buffalo, NY	1.68	1.57	1.44
Wilkes Barre, PA	1.22	1.01	0.95
Harrisburg, PA	1.22	0.61	0.84
Pittsburgh, PA	1.32	1.17	1.76
Baltimore, MD	1.88	3.13	3.04
Richmond, VA	1.33	2.57	2.60
Beckley, WV	1.10	0.77	0.76
Knoxville, TN	1.48	1.13	1.02
Nashville, TN	0.34	0.74	0.85
Louisville, KY	2.45	1.74	1.54
Cincinnati, OH	2.61	2.00	1.87
Columbus, OH	1.97	1.43	1.34
Cleveland, OH	2.96	2.95	2.74
Detroit, MI	2.42	4.58	4.20
Milwaukee, WI	1.32	1.72	1.57
Chicago, IL	3.03	4.32	4.06
Indianapolis, IN	1.77	2.10	2.06
St. Louis, MO	3.21	2.76	2.45
Memphis, TN	1.34	1.44	1.31
Jackson, MS	0.86	0.84	0.79
New Orleans, LA	1.56	1.19	1.05
Montgomery, AL	1.17	2.05	2.00
Atlanta, GA	0.48	1.90	2.46
Fort Jackson, SC	1.13	2.39	2.10
Jacksonville, FL	3.90	2.75	2.38
Miami, FL	2.04	2.79	2.59
Charlotte, NC	1.67	1.53	1.42
Raleigh, NC	1.54	2.01	1.76
Shreveport, LA	0.92	0.82	0.73
Dallas, TX	1.78	1.25	1.46
Houston, TX	1.99	1.33	1.44

San Antonio, TX	2.68	1.53	1.55
Oklahoma City, OK	0.89	0.82	0.84
Amarillo, TX	0.18	0.10	0.16
Little Rock, AR	1.06	1.07	0.99
Kansas City, MO	0.81	0.61	1.31
Des Moines, IA	0.22	0.42	0.54
Minneapolis, MN	1.25	1.61	1.51
Fargo, ND	0.39	0.29	0.32
Sioux Falls, SD	0.51	0.34	0.33
Omaha, NE	0.99	0.79	0.75
Denver, CO	2.57	1.76	1.63
Albuquerque, NM	0.42	0.51	0.56
El Paso, TX	1.07	0.73	0.79
Phoenix, AZ	2.08	1.98	1.80
Salt Lake City, UT	0.50	0.38	0.34
Butte, MT	0.35	0.27	0.29
Spokane, WA	0.73	0.55	0.53
Boise, ID	0.49	0.38	0.34
Seattle, WA	1.62	1.25	1.09
Portland, OR	0.72	1.00	1.19
Oakland, CA	1.79	3.06	3.06
Fresno, CA	1.36	0.94	0.90
Los Angeles, CA	6.73	4.85	4.69
Honolulu, HI	0.56	0.43	0.47
San Juan, PR	1.55	0.37	0.51

BIBLIOGRAPHY

Bishop, Y., S. Fienberg, and P. Holland, *Discrete Multivariate Analysis,* MIT Press, Cambridge, Massachusetts, 1975.

Brown, Morton B., "Screening Effects in Multidimensional Contingency Tables," *Applied Statistics,* Volume 25, Number 1, 1976, pages 37-46.

Doering, Zahava D., David W. Grissmer, and Jane S. Morse, *1979 DoD Survey of Personnel Entering Military Service: Wave 1 User's Manual and Codebook,* The Rand Corporation, N-1605-MRAL, November 1980.

Doering, Zahava D., David W. Grissmer, and Jane S. Morse, *1979 DoD Survey of Personnel Entering Military Service: Wave 2 User's Manual and Codebook,* The Rand Corporation, N-1606-MRAL, December 1980.

DuMouchel, William H., and Greg J. Duncan, "Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples," *Journal of the American Statistical Association,* Volume 78, Number 383, September 1983, pages 535-543.

Frankel, Martin R., and Harold A. McWilliams, *The Profile of American Youth, Technical Sampling Report,* National Opinion Research Center, October 1981.

Holt, D., T.M.F. Smith, and P. D. Winter, "Regression Analysis of Data from Complex Surveys," *Journal of the Royal Statistical Society,* Series A, Volume 143, 1980, pages 474-487.

Institute for Social Research, *A Panel Study of Income Dynamics: Study Design, Procedures, Available Data,* Volume I, 1972.

Jones, Calvin, Geraldine Mooney, and Harold McWilliams, *High School and Beyond 1980 Senior Cohort First Follow-up (1982): Data File User's Manual,* National Center for Educational Statistics, May 1983.

Maddala, G.S., *Econometrics,* McGraw-Hill, New York, 1977.

Manski, Charles F., and Steven R. Lerman, "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica,* Volume 45, Number 8, November 1977, pages 1977-1988.

Nathan, G., and D. Holt, "The Effect of Survey Design on Regression Analysis," *Journal of the Royal Statistical Society,* Series B, Volume 42, 1980, pages 377-386.

Nerlove, Marc, and S. James Press, *Univariate and Multivariate Log-Linear and Logistic Models,* The Rand Corporation, R-1306-EDA/NIH, December 1973.

Porter, Richard D., "On the Use of Survey Sample Weights in the Linear Model," *Annals of Economic and Social Measurement,* Volume 2, 1973, pages 141-158.

Relles, Daniel A., *Using Weights to Estimate Population Parameters from Survey Records,* The Rand Corporation, N-1136-HUD, May 1981.

Smith, Kent W., "Analyzing Disproportionately Stratified Samples with Computerized Statistical Packages," *Sociological Methods and Research,* Volume 5, Number 2, November 1976, pages 207-230.

END

FILMED

8 84

DTIC